Through a glass darkly:

Research biases that result from wearing 'literate glasses'

A commentary on 'The worries of wearing literate glasses'

By R. Kolinsky and J. Morais

Usha Goswami

Centre for Neuroscience in Education

Department of Psychology

University of Cambridge

Author contact details:

Department of Psychology, Downing Street, Cambridge, CB2 3EB

Tel. 44 1223 333550 Fax 44 1223 333564

ucg10@cam.ac.uk

Abstract

In this comment, I argue that Kolinsky and Morais are correct to highlight the multiple research biases that arise from researchers' lack of awareness that they are wearing 'literate glasses'. Converging evidence can be amassed from developmental psychology and indeed, from the widespread use of the International Phonetic Alphabet. Furthermore, cognitive neuroscience as currently practiced is busy repeating the mistakes made cognitive science. Accordingly, fundamental conclusions about how the brain represents information in research fields such as semantic memory, space, time and language may apply only to the educated and literate adult brain.

In their compelling paper, Kolinsky and Morais point out the numerous unconscious biases that result from researchers wearing ‘literate glasses’. By the term ‘literate glasses’, Kolinsky and Morais mean researchers *assuming* that they are studying basic aspects of cognitive and linguistic function in humans when they are *actually* only studying these functions in adults who can read. This blinkered vision results in strong claims being made about basic aspects of cognitive systems such as reasoning, memory and even emotional processes, claims that disregard any potential effects that learning to read may exert on these systems. In their elegant survey, Kolinsky and Morais draw our attention to numerous examples of scientific blindness to the effects of literacy on the brain, encompassing both theoretical and empirical work. To a child psychologist who studies language acquisition and reading acquisition, this is familiar territory. There are many experimental demonstrations of the profound effects of literacy on basic cognitive functions in developmental psychology, and I will give some striking examples in this commentary.

Moreover, and perhaps more immediately relevant to current research in psychology, cognitive neuroscience is now busy repeating the mistakes made by cognitive psychology. New brain imaging techniques enable detailed investigations of the underlying mechanisms supporting human cognition, analyses that were previously impossible. Yet the neural processes underpinning fundamental human cognitive skills such as speech recognition and the organization of semantic memory are being studied without any recognition that these neural mechanisms are being studied in the ‘end states’ of cognitive systems, in highly trained brains. While the information being documented is still interesting and useful, it is not the case that current cognitive neuroscience is studying the fundamental building blocks of cognitive systems such as language, memory and reasoning. Rather, cognitive neuroscientists are investigating the ‘special case’ of such systems in humans who are skilled readers of print.

Take human language. Developmental analyses profoundly contradict the view that Kolinsky and Morais attribute to Chomsky (1975), that communication is only one function of language and by no means an essential one. Instead, developmental psychology has reached the conclusion that the fundamental building blocks for language acquisition are *ostensive signals* that enable infants to recognise *communicative intent* (Csibra, 2010). This specialised set of ostensive signals, which even new-born infants recognize, are specifically designed to generate the interpretation that the communicator has a communicative intention which is addressed to that recipient. The key ostensive signals are establishing mutual eye contact, speaking in infant-directed speech, mutual contingency (turn-taking) and using the infant's name. This developmental perspective changes the nature of research into mechanisms of language acquisition. For example, it turns out that the non-human species who are best at recognizing communicative intent are dogs (Kaminski, Schulz & Tomasello, 2012). Indeed, one dog has learned over 1000 words, and has also learned some simple syntax (Pilley & Hintzmann, 2014).

Similarly, oral language processing changes profoundly with literacy, even in young children. Before children learn to read, they are as accurate at deciding that spoken words like "boat" and "note" rhyme as deciding that words like "bank" and "tank" rhyme. However, even the first year of literacy training changes young children's rhyme judgements (Goswami, Ziegler & Richardson, 2005). Furthermore, differential effects are already visible after just a year of reading instruction for young learners of English, an opaque orthography, versus German, a transparent orthography with similar phonology to English. Oral language processing in young English readers is affected by incongruent rime spellings (the rime is the vowel and any following consonant phonemes in a syllable, for example 'eam' in dr-eam), while the young German readers are affected by incongruent spellings for individual phonemes (how the vowel is spelled, Goswami et al., 2005). Other examples of spelling effects on oral language processing abound in the developmental literature. Pre-readers judge that words like "pitch" and "rich" have the same number of sounds. Early readers count more sounds in "pitch" (Ehri & Wilce, 1980). Precocious spellers, 5-year-olds who have not yet been taught to read or spell and thus have not learned formal orthographic conventions, spell TRUCK as 'CHRAC' and ASHTRAY as 'ASCHRAY' (Read, 1986). The pre-readers are (correctly) perceiving the aspirated 'T' in the consonant cluster TR as sounding more like the sound made by the letters CH than the sound made by the letter T. As Kolinsky and Morais argue, despite all the evidence to the contrary, the educational world still persists in the idea that phonemes are real units, and too many psycholinguists do not distinguish between phoneme awareness and phonological awareness. The latter is truly misguided, as phonological awareness of units such as syllables and rhymes develops prior to literacy, and indeed shows a similar developmental trajectory across languages, in contrast to the trajectory for developing phoneme awareness (Ziegler & Goswami, 2005). Ziegler and Goswami's *psycholinguistic grain size* theory is based on experimental evidence showing that phonological units at larger grain sizes are accessible to pre-reading children, across all languages, while phonemic awareness depends on literacy tuition.

Experimental child data of the kind produced by Read (1986) and Ehri and Wilce (1980) are profoundly important for our understanding of speech perception and linguistic processing. Taken to a logical extreme, such data suggest that all linguistic analysis that is based on the International Phonetic Alphabet (IPA) is misguided, as the IPA itself is the product of literate brains. Accordingly, whenever a linguist analyses any language output using the IPA, that linguist is wearing 'literate glasses'. In 2007, the experimental linguist Robert Port wrote a paper called 'How are words stored in memory? Beyond phones and phonemes' that I consider to be of fundamental importance to the field, but that seems to be poorly known (Port, 2007). Having spent many years carrying out psychoacoustic studies of speech perception and production, Port argued that speech is not processed by the human brain in terms of letter-like symbolic units, namely the consonant and vowel phonemes utilized by the IPA. Port wrote “It seems intuitively obvious that … when we hear someone say a word like tomato, we .. hear a sequence of consonant and vowel sound units” (page 143). Assembling evidence from a range of areas in cognitive science, Port (2007) showed instead that words and other linguistic patterns are stored in memory as detailed sensory representations that incorporate information about speaker properties such as gender and even about other speaker characteristics such as tempo. In Kolinsky and Morais’ terminology, the IPA is a ‘literate hearing aid’. Port argued that spoken language is not learned as a *symbol system* by children, but as *distributions of patterns* in a high-dimensional space of common acoustic patterns (namely phonology). Port’s paper showed the pervasive effects of wearing ‘literate glasses’ in the fields of linguistics and psychology, and indeed he argued that accepting the biases imposed by literacy required a major reconsideration of the goals of these fields. Further, Port argued that recognizing this bias also required a major consideration of what symbols really are, since spoken words are not symbols in the same way that written words are, as Kolinksy and Morais also argue. According to Port (2007), recognizing the role of ‘literate glasses’ necessitates reconceptualising the role that symbolic language occupies in human cognition and in the development of civilization. Port’s paper should be required reading for those convinced by Kolinsky and Morais’ fine piece, as he brings additional relevant evidence to support the cultural points that they make about the effects of ‘literate glasses’ on power relationships, embodied cognition and ‘primitive’ languages.

Turning to cognitive neuroscience, major advances in our understanding of the mechanics of linguistic processing have been made in the last decade, particularly concerning the functional role of neuroelectric oscillations and how they are entrained by speech input

(Giraud & Poeppel, 2012, for review). This branch of cognitive neuroscience has supported Port (2007)'s intuitions, showing that the brain indeed represents the sensory information in the speech signal in rich and concrete detail. Nevertheless, the field is already dogged by wearing 'literate glasses'. I will give 3 examples here, regarding speech perception, speech production, and phonological processing.

Seminal papers in auditory neuroscience by authors such as Poeppel, Giraud, Simon, Gross, Chang and their colleagues has been transformational regarding our understanding of how the speech signal is encoded by the brain (e.g. Luo & Poeppel, 2007; Pasley et al., 2012; Gross et al., 2013; Ding et al., 2014). Detailed studies of the entrainment of neuroelectric oscillations at different temporal rates (delta [1 – 3 Hz, 1 – 3 times per second], theta [4 – 8 Hz], beta [15 – 30 Hz] and gamma [> 30 Hz, rates from Poeppel [2014]) have shown that the quasi-rhythmic alternation of large cell networks between electrical excitation and inhibition is one mechanism for encoding amplitude modulation patterns at similar rates in the speech signal. When speech is heard, the cell networks automatically re-align the timing of their endogenous oscillations so that peaks in excitability systematically track peaks in amplitude modulation. When this alignment process is accurate, then speech becomes intelligible. The cell networks use rise times in amplitude (rates of change of increases in signal intensity) as a basis for this phase re-alignment process. For example, if rise times at the theta rate are removed from natural speech, removing this acoustic trigger for phase re-alignment of theta networks, then speech becomes unintelligible (Doelling et al., 2014). If simple clicks are inserted into the signal at the same temporal points, then speech becomes intelligible again. This demonstration, along with many studies showing that the peak in modulation in speech energy is around 4 – 5 Hz across languages, has led to the belief that theta is the 'master oscillator', playing a core role in successful speech processing across languages.

The reader can probably guess the role of ‘literate glasses’ in reaching this apparently cross-cultural conclusion. All studies to date involve speech perception by highly literate adults (typically university students). Theta may not be the master oscillator for pre-literate individuals, since studies of neuroelectric oscillations in children show that theta entrainment improves with reading acquisition (Power et al., 2012). Further, studies of speech directed to infants shows that the modulation peak is in the delta band (1 – 3 Hz). Infant-directed speech (IDS) was compared with adult-directed speech (ADS) by recording mothers speaking either to their infants or to an experimenter (Leong et al., 2017). Not only did the modulation peak differ for IDS versus ADS, so did the *phase alignment* of different bands of amplitude modulation. For IDS, the slower bands of amplitude modulations, at delta and theta rates, showed significantly greater phase alignment than in ADS. By contrast, the faster bands of amplitude modulations, at theta and gamma rates, showed significantly greater phase alignment in ADS than in IDS (Leong et al., 2017). Perhaps theta is not the master oscillator until the brain has learned to read.

Similar effects of ‘literate glasses’ can be shown in speech production. The speech produced by highly literate adults and by adults who have never learned to read is different (Aruajo et al., 2018). Aruajo and his colleagues compared the conversational speech of Portuguese adults who had never learned to read because of lack of access to schooling with the conversational speech of Portuguese adults who had many years of literacy education and also with low literate Portuguese adults, who had received up to 4 years of literacy instruction. Aruajo et al. reported significant differences between the conversational speech of the illiterate adults compared to the low and high literates. The illiterate adults produced spontaneous speech that showed significantly less phase alignment between all bands of amplitude modulations. Both delta-theta phase alignment and theta-gamma phase alignment was reduced for the illiterate adults, but just 4 years of literacy tuition changed this pattern to match that of the high literates. The amount of theta-band energy in conversational speech was related to years of literacy tuition. Accordingly, the conclusions taken from current cognitive neuroscience studies of speech processing may not apply to all humans. They may simply apply to the special sub-set of human adults who are literate.

Finally, this new cognitive neuroscience perspective on human speech processing makes theoretical claims about the timescales involved in phonological processing. These claims are interesting and important, and I personally have found them very productive in thinking about (for example) why phonological awareness follows the developmental trajectory documented by psycholinguistic grain size theory (Ziegler & Goswami, 2005). Yet the claims have not panned out as I initially expected in my developmental studies (Leong & Goswami, 2017; Flanagan & Goswami, 2018). The original claim (Poeppel et al., 2008) was that neuroelectric oscillations at the theta rate were relevant to parsing syllables from speech, while neuroelectric oscillations at the gamma rate were relevant to parsing phonemes. Theta oscillations contain on average 5 cycles per second, giving a syllable every 200 ms, which accords with studies of speech production (at least, studies of speech production by literate adults). Gamma oscillations contain on average 40 cycles per second, giving a phoneme approximately every 25 ms, which again accords with speech production by literate adults. Surprisingly, however, analysis of the items used in phonological awareness tasks for *children* did not fit easily into this model. In particular, phoneme awareness tasks did not seem systematically to vary gamma band information. This was found by modelling the amplitude modulation structure of items used in tasks such as phoneme deletion ("Say 'hift' without the /f/; Say 'crots' without the /t/"; Flanagan & Goswami, 2018). Flanagan and Goswami modelled both the items, like "hift", and the targets, like "hit", and then looked for the acoustic factors that systematically differentiated items from targets from an amplitude modulation perspective. They found that changes in the magnitude of synchronization between the slower timescales, delta-band and theta-band amplitude modulations, were the only systematic acoustic cue to phoneme deletion. Changes in the magnitude of synchronization between the faster timescales, theta-band and low gamma-band amplitude modulations, showed no systematic relations with phoneme deletion.

Flanagan and Goswami (2018) also modelled a child morphology task, plural elicitation (“This is a wug. Now there is another one. There are two of them. There are two ?? [wugs]”; Berko, 1958). As pluralization in English involves adding a morpheme (s) that is a single phoneme, this task can be thought of as phoneme addition. Again, the modelling of the amplitude modulation structure of items (like ‘wug’) and correct responses (like ‘wugs’) showed that changes in the magnitude of synchronization between the delta-band and theta-band amplitude modulations were the only systematic acoustic cue to inflectional morphology. Faster gamma band modulations did not play any systematic role in distinguishing items and responses. Accordingly, while rapid temporal information in speech is clearly relevant to distinguishing phonetic features such as formant transitions, it may be slower amplitude information that is relevant to the acquisition of distinctions that we think of as phonemes in human language. Further research is needed, but it may turn out that the brain uses statistical structures in speech, such as amplitude modulation hierarchies (Leong & Goswami, 2015), that were not imagined by cognitive neuroscientists who were wearing their ‘literate glasses’. Kolinsky and Morais are correct to argue that wearing ‘literate glasses’ leads us to underestimate the contribution of literacy to cognition. As they state, ignoring the fact that we are all wearing ‘literate hearing aids’ leads to a biased view of the basic parameters underlying human language and human cognition.

Bibliography

Araujo, J., Flanagan, S., Castro-Caldas, A., & Goswami, U. (2018). The temporal modulation structure of illiterate versus literate adult speech. *PLoS One, 13 (10*), e0205224.

Berko, J. (1958). The child's learning of English morphology. *Word*. *14*, 150-177.

Chomsky, N. (1975). Reflections on Language. Glasgow: Fontana/Collins.

Csibra, G. (2010). Recognising communicative intentions in infancy. *Mind & Language,* 25, 121 – 168.

Ding, N., Chatterjee, M., & Simon, J.Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *Neuroimage, 88*, 41-6.

Doelling, K.B., Arnal, L.H., Ghitza, O., & Poeppel, D. (2014) Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage. 85*, 761–768.

Ehri, L.C. & Wilce, L.S. (1980). The influence of orthography on readers' conceptualisation of the phonemic structure of words. *Applied Psycholinguistics, 1,* 371-385.

Flanagan, S.A. & Goswami, U. (2018). The role of phase synchronisation between low frequency amplitude modulations in child phonology and morphology speech tasks. *Journal of the Acoustical Society of America.* 143(3), 1366 – 1375.

Giraud, A.L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience, 15***,** 511-517.

Goswami, U., Ziegler, J., & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology, 92*, 345-365.

Gross J., Hoogenboom N., Thut G., Schyns P., Panzeri S., Belin P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology, 11 (12),* e1001752.

Kaminiski, J., Schulz, L., & Tomasello, M. (2012). How dogs know when communication is intended for them. *Developmental Science, 15 (2),* 222-232.

Leong ,V., and Goswami, U. (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PLoS One, 10 (12),* e0144411.

Leong, V., Kalashnikova, M., Burnham, D. & Goswami, U. (2017). The temporal modulation structure of infant-directed speech. *Open Mind, 1,* (2), p.78-90.

Luo, H., and Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron, 54*, 1001-10.

Pasley, B.N., David, S.V., Mesgarani, N., Flinker, A., Shamma, S.A., Krone, N.E., Knight, R.T. & Chang, E.F. (2012). Reconstructing speech from human auditory cortex. *PLoS Biology, 10 (1*), e1001251.

Pilley, J.W. & Hinzmann, H. (2014). *Chaser: Unlocking the genius of the dog who knows 1000 words*. Oneworld Publications.

Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology, 28c*, 142-149.

Poeppel, D., Idsardi, W. J., & Van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B, 363, 1071-1086.*

Port, R. (2007). How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology, 25*, 143-70.

Power, A.J., Mead, N., Barnes, L., & Goswami, U. (2012). Neural entrainment to rhythmically-presented auditory, visual and audio-visual speech in children. *Frontiers in Psychology, 3*, 216.

Read, C. (1986). *Children's creative spelling*. London, UK: Routledge Kegan Paul.

Ziegler, J.C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3-29.